



CREATING SCALABLE AI INFRASTRUCTURE

Wei Li
VP/GM, Machine Learning Performance

Oct 2019

INTEL AI HARDWARE FROM EDGE TO DATA CENTER



DEVICE



INTELLIGENT EDGE



MULTI-CLOUD

OPTIMIZED FRAMEWORKS & SOFTWARE

CPU



GPU



FPGA



ASIC



WORKLOAD BREADTH

AI SPECIALIZATION

Multi-Purpose
Foundation for AI

Data-Parallel Media,
Graphics, HPC & AI

Multi-Function & Real-time
Deep Learning Inference

Deep Learning
Inference

Deep Learning
Training

Media & Vision
DL Inference at
the Edge

Visit:

www.intel.ai/technology

All products, computer systems, dates, and figures are preliminary based on current expectations, and are subject to change without notice.
1 Unified software stack development in progress DL=Deep Learning

INTEL OPTIMIZED AI SOFTWARE

**DATA SCIENTISTS and
APP DEVELOPERS**



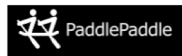
Caffe *

mxnet *



Caffe2 *

PyTorch *



**OpenVINO
Inference
Engine**

Intel® Deep Neural Network Library (DNNL)

Computational Primitives

**Intel® Machine Learning
Scaling Library (Intel® MLSL)**

Communication Primitives

SEE ALSO: Machine Learning Libraries for Python (Scikit-learn, Pandas, NumPy), R (Cart, randomForest, e1071), Distributed (MLlib on Spark, Mahout)

* Other names and brands may be claimed as the property of others.

ACCELERATING CONVERGENCE

LEADERSHIP PERFORMANCE FOR DATA LEVEL PARALLEL HPC & AI WORKLOADS



7NM PROCESS TECHNOLOGY

EMIB (2D) AND FOVEROS (3D) TECHNOLOGY

POWERING AURORA
DELIVERED IN 2021



**PARTNER
WITH INTEL
TO ACCELERATE
YOUR AI
JOURNEY**

INTEL® AI WITH SCALE



SPEED UP DEVELOPMENT
with open AI software



DEPLOY AI ANYWHERE
with unprecedented HW choice



SCALE WITH CONFIDENCE
on the engine for IT & cloud

LEARN MORE AT WWW.INTEL.AI